

✓

· ID : CA13_BOVIN STANDARD PRT; 1049 AA.
AC P04258;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ~~collagen alpha 1(III) chain.~~
GN COL3A1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-242.
RX MEDLINE=80026026; PubMed=488906;
RA Fietzek P.P., Allmann H., Rauterberg J., Henkel W., Wachter E.,
RA Kuhn K.;
RT "The covalent structure of calf skin type III collagen. I. The amino
acid sequence of the amino terminal region of the alpha 1(III) chain
(positions 1-222).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:809-820(1979).
RN [2]
RP SEQUENCE OF 243-422.
RX MEDLINE=80026027; PubMed=488907;
RA Dewes H., Fietzek P.P., Kuhn K.;
RT "The covalent structure of calf skin type III collagen. II. The amino
acid sequence of the cyanogen bromide peptide alpha 1(III)CB1,8,10,2
(positions 223-402).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:821-832(1979).
RN [3]
RP SEQUENCE OF 423-571.
RX MEDLINE=80026028; PubMed=488908;
RA Bentz H., Fietzek P.P., Kuhn K.;
RT "The covalent structure of calf skin type III collagen. III. The amino
acid sequence of the cyanogen bromide peptide alpha 1(III)CB4
(positions 403-551).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:833-840(1979).
RN [4]
RP SEQUENCE OF 572-808.
RX MEDLINE=80026029; PubMed=488909;
RA Lang H., Glanville R.W., Fietzek P.P., Kuhn K.;
RT "The covalent structure of calf skin type III collagen. IV. The amino
acid sequence of the cyanogen bromide peptide alpha 1(III)CB5
(positions 552-788).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:841-850(1979).
RN [5]
RP SEQUENCE OF 809-947.
RX MEDLINE=80026030; PubMed=488910;
RA Dewes H., Fietzek P.P., Kuhn K.;
RT "The covalent structure of calf skin type III collagen. V. The amino
acid sequence of the cyanogen bromide peptide alpha 1(III)CB9A
(position 789-927).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:851-860(1979).
RN [6]
RP SEQUENCE OF 948-1049.
RX MEDLINE=80026031; PubMed=488911;
RA Allmann H., Fietzek P.P., Glanville R.W., Kuhn K.;
RT "The covalent structure of calf skin type III collagen. VI. The amino
acid sequence of the carboxyterminal cyanogen bromide peptide alpha
1(III)CB9B (positions 928-1028).";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:861-868(1979).
CC -!- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
CC -!- ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE

CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 DR PIR; A02862; CGB075.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01391; Collagen; 17.
 DR ProDom; PD000007; Collagen; 3.
 DR PROSITE; PS01208; VWF_C_1; PARTIAL.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen.
 FT DOMAIN 1 14 NONHELICAL REGION (N-TERMINAL).
 FT DOMAIN 15 1040 TRIPLE-HELICAL REGION.
 FT DOMAIN 1041 1049 NONHELICAL REGION (C-TERMINAL).
 FT MOD_RES 95 95 HYDROXYLATION.
 FT MOD_RES 107 107 HYDROXYLATION.
 FT MOD_RES 119 119 HYDROXYLATION.
 FT MOD_RES 938 938 HYDROXYLATION.
 FT MOD_RES 950 950 HYDROXYLATION.
 FT CARBOHYD 107 107 O-LINKED (GAL...).
 FT CARBOHYD 950 950 O-LINKED (GAL...).
 FT DISULFID 1040 1040 INTERCHAIN.
 FT DISULFID 1041 1041 INTERCHAIN.
 SQ SEQUENCE 1049 AA; 93651 MW; 8EEC33D1C66EC9A3 CRC64;
 EYEAYDVKSG VAGGGIAGYGP GPAGPAGPQQP PPCTSGHPGA PGAPGYQGPP GEPGQAGPAG
 PPGPPGAIGP SGKDGESGRP GRPQPRGFPG PPGMKGPAGM GFDPGMKGHR GFDRGRNGEKG
 EPGAPGLKG NGVPGEDGAP GPMGPGRGPG ERGRPGLPG AGARNDGAR GSDCQGPQPG
 PPGTAGFPGSG PGAKGEVGPAG GSPGSSGAPG QRGEPPQOH AGAPGPQGP GSDGSPGKG
 EMGPAGIPGA PGLIGARGPPP GPPGTTNGVPG QRGAAGEPCK NGAKGDPGP GERGEAGSPG
 IAGPKGEDGPN DGSPEGPAGN GLPGAAAGERG VPGFRRGPAGA NGLPGEKGPP GDRGGPQGPAG
 PRGVAGEPGR NGLPGPGLR GIPGSPGPPG SNKGPQPPGS QGETGRGPQG GSPGPQGPQG
 VMGFPGPKGDN DGA PGKNGER GGPQGPQGPQG PAGKNGETGP OGPPGPQGP GDKGDTGPQG
 PQGLQGLPQGT SGPPGENGKP GEPPKGKEAG APGIPGKGKD SGAPGERGP GAGGPPGPQG
 GAGGPQGPQG KGAAGFPGPQ GSAGTGPGLQG MPGERCGPQG PCPKGDKGPP GSSGVDPQGP
 KGDRGPQPTGQ IGPPGPQGPQG GDKGESGAPG VPGIAGPRGG PGERGEQGPQ GPAGFPQGPQ
 QNGEPQAKGR RGAPEKGEG GPPGAAGPAG GSPGAGPQGPQ QGVKGERGP GPPGAAFPQGP
 GRGPPGPQGPQ NGNPQGPQGSS GAPGKDKGPQ PGPSNQAPGS PG1SGPKDQ GPPGERGPAG
 PQGPQGPQGPQ LGIAQLTQAR GLAGPQGPMPG ARGSPQGPQGI KGENGKPGPS QGNGERGPQGP
 PQGLPGLAGT AGEPGRDGNP GSDGLPGRDG APGAKGDRGE NGSPGPQGP GHPGPPGPVQ
 PAGKSGDRGEG TGPAGPSGAP GPAGSQRGPQ PGQPRGDKE TGERGAMOIK GHRGFPQGPQ
 APGSPGPAGH QGAVGSGPQGA GPRGPVGPQ PGPKDQASGH PGPIGPQGPQ GNRGERGSEG
 SPGPQGPQGP PGPPGAPGPQ CGAGGVAI

//

ID CA13_HUMAN STANDARD; PRT; 1466 AA.
AC P02461; Q15112.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(III) chain precursor.
GN COL3A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN FIBROBLAST;
RX MEDLINE=89350838; PubMed=2764886;
RA Ala-Kokko L., Kontusaari S., Baldwin C.T., Kuvianiemi H.,
RA Prockop D.J.;
RT "Structure of cDNA clones coding for the entire prepro alpha 1 (III)
RT chain of human type III procollagen. Differences in protein structure
RT from type I procollagen and conservation of codon preferences.";
RL Bicchem. J. 260:509-516(1989).
RN [2]
RP SEQUENCE OF 149-1225 FROM N.A.
RX MEDLINE=89386015; PubMed=2780304;
RA Janeczko R.A., Ramirez F.;
RT "Nucleotide and amino acid sequences of the entire human alpha 1 (III)
RT collagen.";
RL Nucleic Acids Res. 17:6742-6742(1989).
RN [3]
RP SEQUENCE OF 168-398.
RX MEDLINE=77134724; PubMed=557335;
RA Seyer J.M., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of cyanogen
RT bromide peptides from the amino-terminal segment of type III collagen
RT of human liver.";
RL Biochemistry 16:1158-1164(1977).
RN [4]
RP REVISIONS.
RA Seyer J.M.;
RL Submitted (DEC-1977) to the PIR data bank.
RN [5]
RP SEQUENCE OF 399-727.
RX MEDLINE=79000343; PubMed=687591;
RA Seyer J.M., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of five
RT consecutive CNBr peptides from type III collagen of human liver.";
RL Biochemistry 17:3404-3411(1978).
RN [6]
RP SEQUENCE OF 728-964.
RX MEDLINE=80198282; PubMed=6246925;
RA Seyer J.M., Mainardi C., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of alpha 1 (III)-
RT CB5 from type III collagen of human liver.";
RL Biochemistry 19:1583-1589(1980).
RN [7]
RP SEQUENCE OF 950-1466 FROM N.A.
RX MEDLINE=88189827; PubMed=3357782;
RA Mankoo B.S., Dalglish R.;
RT "Human pro alpha 1(III) collagen: cDNA sequence for the 3' end.";
RL Nucleic Acids Res. 16:2337-2337(1988).
RN [8]
RP REVISION TO 1184.
RX MEDLINE=89098346; PubMed=32211760;
RA Molyneux K., Dalglish R.;
RT "Human type III collagen 'variant' is a cDNA cloning artefact.";

FT VARIANT 1170 1170 G -> V (IN EDS-IV).
 FT VARIANT 1173 1173 /FTID=VAR_011157.
 FT VARIANT 1173 1173 G -> E (IN EDS-IV).
 FT VARIANT 1173 1173 /FTID=VAR_001801.
 FT VARIANT 1176 1176 G -> R (IN GROTTON TYPE ACROGERIA).
 FT VARIANT 1176 1176 /FTID=VAR_011158.
 FT VARIANT 1179 1179 G -> V (IN EDS).
 FT VARIANT 1179 1179 /FTID=VAR_001802.
 FT VARIANT 1182 1182 G -> R (IN EDS-IV).
 FT VARIANT 1182 1182 /FTID=VAR_011159.
 FT VARIANT 1185 1185 G -> E (IN EDS).
 FT VARIANT 1185 1185 /FTID=VAR_001803.
 FT VARIANT 1185 1185 G -> D (IN EDS-IV; SEVERE VARIANT).
 FT VARIANT 1185 1185 /FTID=VAR_001804.
 FT VARIANT 1188 1188 G -> V (IN EDS-IV).
 FT VARIANT 1188 1188 /FTID=VAR_001805.
 FT VARIANT 1188 1188 G -> E (IN EDS-IV; SEVERE VARIANT).
 FT VARIANT 1188 1188 /FTID=VAR_001806.
 FT VARIANT 1188 1188 G -> R (IN EDS-IV).
 FT VARIANT 1188 1188 /FTID=VAR_001807.
 FT CONFLICT 163 163 G -> GG (IN REF. 2).
 FT CONFLICT 168 168 G -> V (IN REF. 4).
 FT CONFLICT 226 228 MISSING (IN REF. 4).
 FT CONFLICT 241 241 E -> D (IN REF. 2).
 FT CONFLICT 278 278 T -> A (IN REF. 4).
 FT CONFLICT 293 295 NGA -> DGS (IN REF. 4).
 FT CONFLICT 472 472 E -> D (IN REF. 2).
 FT CONFLICT 488 490 PGF -> LGS (IN REF. 2).
 FT CONFLICT 614 614 T -> Y (IN REF. 2).
 FT CONFLICT 635 635 P -> R (IN REF. 2).
 FT CONFLICT 664 664 D -> E (IN REF. 2).
 FT CONFLICT 676 676 D -> N (IN REF. 5).
 FT CONFLICT 896 896 S -> A (IN REF. 6).
 FT CONFLICT 980 980 S -> A (IN REF. 10).
 FT CONFLICT 985 989 ANGLS -> PSGQN (IN REF. 10).
 FT CONFLICT 1019 1019 D -> Y (IN REF. 7).
 FT CONFLICT 1097 1097 T -> P (IN REF. 10).
 FT CONFLICT 1153 1154 TS -> AT (IN REF. 10).
 FT CONFLICT 1156 1156 H -> P (IN REF. 9).
 FT CONFLICT 1156 1156 H -> S (IN REF. 10).
 FT CONFLICT 1184 1184 P -> S (IN REF. 2).
 FT CONFLICT 1203 1203 A -> P (IN REF. 2).
 FT CONFLICT 1210 1210 G -> A (IN REF. 2).
 FT CONFLICT 1241 1241 V -> A (IN REF. 7 AND 9).
 FT CONFLICT 1357 1357 L -> P (IN REF. 11).
 SQ SEQUENCE 1466 AA; 138556 MM; 990E3ACB33A10847 CRC64;
 MMSFVQKGSW LLLLALHPTI ILAQOBAVEG GCSHLLGOSYA DRDWVKEPEPC QICVCDGSV
 LCDDII CDDQ ELDCPNPEIP FGECACVCPQ PPTAPTRPPN GCGPQPKPGD PGPGIPIGRN
 GDFGIPQPG SPGSPGPGI CESCPTGPQN YSPQYDSDV KSGVAVGGLA GYPGPAGPG
 PPGPGTSHG PGSPGSPGQY GPPGEPQAG QSPGPQPGN IGSPGPAGKD GESGRPGRPG
 ERGLPGPQGI KGPAGIPGPF GMKGHRGFD RNEKEGETGA PGLKGENGLP GENGAPGPMG
 PRGAPGERER PGLPGAAAGR GNDGARGSDG QPGPQPGPQG AFGPGSPGAK GEVGPAGSPG
 SNGAPGPGRE PGPGQHAGAQ GPPGPPGNG SPGKGEMGAG AIGPAGPGLM GARGPPGPG
 ANAGPGLRGG AEGPQKNGAK GEPPGPERG EAGIPGVGGA KGEDGKDGP GEPGANGLPG
 AAGERGAPFG RGPAGPONGIP GEKGPAGPERG APGPAGPRGA AEGPGRDGVP GCPOMRQWPG
 SPGPGPSDGG PGPGPSQGES GRGPQPGPSG PRQGPQVMGP PGPKQNDGAP GKNGERGPG
 GPQPGPQPKG NGETGPQGPQ GPTGPQGDKG DTGPQPGQL QGLPGTGGPP GENKGPEPG
 PKDQDAGAPGA PGKGKDAGAP GERGPQPLAG APGLRGGAGP PGPEGGKGAA GPPGPPQAG
 TPLGLQMPGE RGGLGSPGPK GDKGPQPGPG ADGVPGKDGP RGPPTGPQGPQ PGAGQPGDKG
 EGGAPGLPGI AGRGPSPGER GETGPQPGAG FPGAGPQNGEN PGKGERGAP GEKGEGGPPG
 VAGPGPQSGP AGPQPGPKVQ GERSGPQPG AAGFGPQGRL PGPGNSNGP PGPGPSGSPG
 KDGPGPAGN TGAPGSPGVs GPKQDAGQPG EKGSQPGQGP PGAPGFLCIA GITGARCLAG
 PGMPGPGRGS PGPGQVKGES GKPGANGLGS ERGPQPGQGL PGLAGTAGEP GRDGNPSDQ
 LPGRDGPQGK KGDGRGENSP GAPGAPGHG PPGPVGPAKG SGDREGESGPQ GPAGAPGPAG
 SRGAPGPQGP RGDKGETGER GAAGIKHHRG FPQNPQGPQG PGPAGQGQAI GSPGPGAPRPG

2 of 3

1 of 13

• P V G P S G P P G K D G T S G H P G P I G P P G P R G N R G E R G S E G S P G H P G Q P G P P G P P G A P G P C C G G V G A A A I A G I G G E K A G G F A P Y Y G D E P M D F K I N T D E I M T S L K S V N G Q I E S L I S P D G S R K N P A R • N C R D L K F C H T E L K S G E Y W D P N Q C C K L D A I K V F C N M E T G E T C I S A N P L N V P R K H W T D S S A E K K H V W F G E S M D G F Q F S Y G N P E L P E D V L D V Q L A F L R L L S S R A S Q N I T Y H C K N S I A Y M D Q A S G N V K K A L K L M G S N E G E F K A E G N S K F T Y T V L E D G C T K H T G E W S K T V F E Y R T R K A V R L P I V D I A P Y D I G G P D Q E F G V D V G P V C F L

//